Epidemiology and Infection

Systematic review of risk prediction scores for surgical site infection or periprosthetic joint infection following joint arthroplasty

S. K. Kunutsor, M. R. Whitehouse, A. W. Blom, A. D. Beswick

Supplementary Material

Supplementary Material 1	PRISMA checklist
Supplementary Material 2	Materials and methods
Supplementary Material 3	Literature search strategy
Supplementary Material 4	Quality assessment of included risk scores using the PROBAST tool

Supplementary Material 1. PRISMA Checklist

Section/topic	Item No	Checklist item	Reported on page No
Title			
Title	1	Identify the report as a systematic review, meta-analysis, or both	1
Abstract			
Structured summary	2	Provide a structured summary including, as applicable, background, objectives, data sources, study eligibility criteria, participants, interventions, study appraisal and synthesis methods, results, limitations, conclusions and implications of key findings, systematic review registration number	2
Introduction			
Rationale	3	Describe the rationale for the review in the context of what is already known	4-5
Objectives	4	Provide an explicit statement of questions being addressed with reference to participants, interventions, comparisons, outcomes, and study design (PICOS)	5
Methods			
Protocol and registration	5	Indicate if a review protocol exists, if and where it can be accessed (such as web address), and, if available, provide registration information including registration number	Materials and Methods
Eligibility criteria	6	Specify study characteristics (such as PICOS, length of follow-up) and report characteristics (such as years considered, language, publication status) used as criteria for eligibility, giving rationale	Materials and Methods
Information sources	7	Describe all information sources (such as databases with dates of coverage, contact with study authors to identify additional studies) in the search and date last searched	Materials and Methods
Search	8	Present full electronic search strategy for at least one database, including any limits used, such that it could be repeated	Supplementary Material 3
Study selection	9	State the process for selecting studies (that is, screening, eligibility, included in systematic review, and, if applicable, included in the meta-analysis)	Materials and Methods
Data collection process	10	Describe method of data extraction from reports (such as piloted forms, independently, in duplicate) and any processes for obtaining and confirming data from investigators	Materials and Methods
Data items	11	List and define all variables for which data were sought (such as PICOS, funding sources) and any assumptions and simplifications made	Materials and Methods
Risk of bias in individual studies	12	Describe methods used for assessing risk of bias of individual studies (including specification of whether this was done at the study or outcome level), and how this information is to be used in any data synthesis	Materials and Methods
Summary measures	13	State the principal summary measures (such as risk ratio, difference in means).	Not applicable
Synthesis of results	14	Describe the methods of handling data and combining results of studies, if done, including measures of consistency (such as I^2 statistic) for each meta-analysis	Not applicable
Risk of bias across studies	15	Specify any assessment of risk of bias that may affect the cumulative evidence (such as publication bias, selective reporting within studies)	Materials and Methods
Additional analyses	16	Describe methods of additional analyses (such as sensitivity or subgroup analyses, meta-regression), if done, indicating which were pre-specified	Not applicable
Results			
Study selection	17	Give numbers of studies screened, assessed for eligibility, and included in the review, with reasons for exclusions at each stage, ideally with a flow diagram	Results and Figure 1
Study characteristics	18	For each study, present characteristics for which data were extracted (such as study size, PICOS, follow-up period) and provide the citations	Tables 1 and 2
Risk of bias within studies	19	Present data on risk of bias of each study and, if available, any outcome-level assessment (see item 12).	Results; Supplementary Material 4
Results of individual studies	20	For all outcomes considered (benefits or harms), present for each study (a) simple summary data for each intervention group and (b) effect estimates and confidence intervals, ideally with a forest plot	Not applicable
Synthesis of results	21	Present results of each meta-analysis done, including confidence intervals and measures of consistency	Not applicable
Risk of bias across studies	22	Present results of any assessment of risk of bias across studies (see item 15)	Not applicable
Additional analysis	23	Give results of additional analyses, if done (such as sensitivity or subgroup analyses, meta-regression) (see item 16)	Not applicable
Discussion			
Summary of evidence	24	Summarise the main findings including the strength of evidence for each main outcome; consider their relevance to key groups (such as health care providers, users, and policy makers)	Discussion
Limitations	25	Discuss limitations at study and outcome level (such as risk of bias), and at review level (such as incomplete retrieval of identified research, reporting bias)	Discussion
Conclusions	26	Provide a general interpretation of the results in the context of other evidence, and implications for future research	Discussion
Funding			
Funding	27	Describe sources of funding for the systematic review and other support (such as supply of data) and role of funders for the systematic review	Funding Section

Supplementary Material 2. Materials and methods

Data sources and search strategy

We searched MEDLINE, EMBASE, Web of Science, and The Cochrane Library electronic databases up to September 30, 2016. The publicly available trial registers ClinicalTrials.gov, UK Clinical Research Network Study Portfolio Database (UKCRN), and the WHO International Clinical Trials Registry Platform were also searched. The search strategy combined free and MeSH search terms and combination of key words relating to risk prediction (e.g., "predict", "risk score", "sensitivity"), PJI (e.g., "periprosthetic joint infection", "deep infection", "surgical site infection"), and joint replacement (e.g., "hip replacement", "knee replacement", "hip arthroplasty", "knee arthroplasty"). No restrictions were placed on publication dates and only articles published in English were considered. Reference lists of retrieved articles and relevant review articles identified on the topic were manually scanned for all relevant additional studies. **Supplementary Material 3** provides details of the search strategy.

Eligibility criteria

Studies were eligible for inclusion if they met the following inclusion criteria: (i) they were population-based studies (prospective or retrospective case control, prospective cohort, retrospective cohort, case-cohort, or nested-case control) that developed or validated a risk prediction score for PJI or an update on a previously developed model; (ii) the endpoint was PJI or reported as a surgical site infection in a longitudinal design; (iii) included patients ≥ 18 years who had been followed up after joint arthroplasty; and (iv) RCTs that assessed the clinical effectiveness of a risk score in an intervention group compared to usual care in a control group. We excluded (i) cross-sectional and clinical case studies; (ii) studies with risk assessment scores containing less than two variables or risk factors; and (iii) studies only reporting measures of associations between an exposure and the risk of PJI.

Study definitions

The performance of a risk prediction score is commonly assessed using measures which include calibration, discrimination, and reclassification.[1] Calibration as measured by the goodness-of-fit statistic, is the ability to correctly estimate the risk of a future event. Discrimination is the ability of the risk score to separate individuals at higher risk from those at lower risk; it is assessed using the area under the receiver-operating characteristic (ROC) curve or C statistic (Harrell's C-index[2]).[3] A C-index of 0.5 represents no improvement over what would be expected by chance, whiles a C-index of 1 implies perfect discrimination.[4] Reclassification or risk stratification, which is assessed using the net-reclassification-improvement (NRI)[1, 5] and the integrated-discrimination-improvement (IDI),[1] is the ability of the risk score to appropriately reclassify patients into clinically relevant subgroups.

Data extraction and quality assessment

Two authors (S.K.K., A.D.B) independently extracted data and a consensus was reached in case of any inconsistency with involvement of a third (M.R.W). We used a predesigned standardized data abstraction form to obtain relevant information. These included information on first author's name, study publication date, country, study design, type of population, statistical model or methods employed, sample size, length of follow-up, infection outcomes, components of each prediction model, measures of discrimination, calibration, and/or reclassification, and reported performance comparisons of the model. Details of validation (internal and/or external) performance statistics were also extracted. Additionally, in the case of publications reporting on multiple risk scores or external validation of previous scores, details of all risk models and their performance indices were extracted and reported separately. The risk of bias (quality) of any study developing or evaluating a risk score was assessed using an early version of the Prediction study Risk Of Bias Assessment Tool (PROBAST), a tool used for assessing risk of bias and applicability of prognostic model studies.[6] Briefly, it uses information on five pre-defined domains namely: participant selection, predictors, outcomes, sample size and patient flow, and analysis. The PROBAST evaluation is used to determine the risk of bias of the risk score (i.e., whether the score is likely to work as intended for the population

of interest), with risk scores classified as low, moderate or high risk of bias. It also assesses the applicability and usability of the risk model.

Data synthesis and analysis

It was planned to synthesize measures of discrimination and calibration if multiple studies were found to have validated the same risk score; however, given the limited number of studies, type of measures reported, and the diversity of the study designs and populations, a formal meta-analysis could not be performed. This was anticipated as it is generally difficult to synthesize the literature on risk prediction due to variability across studies. Effective comparisons could also not be made across studies because of the heterogeneity of the data and the variable methodologies adopted. The characteristics of each study and risk models or scores were summarized in tables. A narrative synthesis was performed according to previously reported quality criteria for risk scores such as usability (10 or fewer components), good calibration, discriminative ability (> 0.70), generalizability (externally validated), and clinical effectiveness.[7, 8]

References

(1) **Pencina MJ, et al.** Evaluating the added predictive ability of a new marker: from area under the ROC curve to reclassification and beyond. *Statistics in Medicine* 2008; **27**(2): 157-172; discussion 207-112.

(2) **Harrell FE, Jr., Lee KL, Mark DB.** Multivariable prognostic models: issues in developing models, evaluating assumptions and adequacy, and measuring and reducing errors. *Statistics in Medicine* 1996; **15**(4): 361-387.

(3) **Cook NR.** Use and misuse of the receiver operating characteristic curve in risk prediction. *Circulation* 2007; **115**(7): 928-935.

(4) **Harrell FE, Jr., et al.** Evaluating the yield of medical tests. *Journal of the American Medical Association* 1982; **247**(18): 2543-2546.

(5) **Pencina MJ, D'Agostino RB, Sr., Steyerberg EW.** Extensions of net reclassification improvement calculations to measure usefulness of new biomarkers. *Statistics in Medicine* 2011; **30**(1): 11-21.

(6) **Anon.** Prediction model study Risk Of Bias Assessment Tool (PROBAST). Available from: http://www.systematic-reviews.com/probast/. Accessed on 12 July, 2016.

(7) **Altman DG, et al.** Prognosis and prognostic research: validating a prognostic model. *British Medical Journal* 2009; **338**: b605.

(8) **Noble D, et al.** Risk models and scores for type 2 diabetes: systematic review. *British Medical Journal* 2011; **343**: d7163.

Supplementary Material 3. Literature search strategy

Relevant studies, published before September 30, 2016 (date last searched), were identified through electronic searches not limited to the English language using MEDLINE, EMBASE, Web of Science, and Cochrane databases. Electronic searches were supplemented by scanning reference lists of articles identified for all relevant studies (including review articles) and by hand searching of relevant journals. The computer-based searches combined search terms related to risk prediction, periprosthetic joint infection, and joint replacement.

1 exp Prosthesis-Related Infections/ or peri-prosthetic joint infection.mp. (9560)
2 prosthetic joint infection.mp. (672)
3 prosthetic infection.mp. (349)
4 peri-prosthetic infection.mp. (46)
5 wound infection.mp. or exp Wound Infection/ (49411)
6 exp Surgical Wound Infection/ or deep infection.mp. (33470)
7 sepsis.mp. or Sepsis/ (100349)
8 surgical site infection.mp. or exp Surgical Wound Infection/ (32804)
9 risk score.mp. or Risk Assessment/ (209689)
10 predict.mp. (246530)
11 score.mp. (380704)
12 diagnostic.mp. (671629)
13 sensitivity.mp. or exp "Sensitivity and Specificity"/ (1060200)
14 specificity.mp. or exp "Sensitivity and Specificity"/ (1056788)
15 exp ROC Curve/ or ROC.mp. (54840)
16 receiver operating characteristic.mp. or exp ROC Curve/ (57559)
17 area under curve.mp. or exp Area Under Curve/ (33411)
18 C statistics.mp. (1047)
19 C-index.mp. (1274)
20 Concordance statistic.mp. (83)
21 prognostic.mp. (208417)
22 algorithm.mp. or exp Algorithms/ (304419)
23 model.mp. (1540191)
24 risk calculator.mp. (405)
25 arthroplasty.mp. or exp Arthroplasty, Replacement/ or exp Arthroplasty, Replacement, Knee/ or exp Arthroplasty/ or exp Arthroplasty, Replacement, Hip/ or exp Arthroplasty, Replacement, Ankle/ or exp Arthroplasty, Replacement, Elbow/ (67099)
26 1 or 2 or 3 or 4 or 5 or 6 or 7 or 8 (158049)
27 9 or 10 or 11 or 12 or 13 or 14 or 15 or 16 or 17 or 18 or 19 or 20 or 21 or 22 or 23 or 24 (4333716)
28 25 and 26 and 27 (1829)
29 limit 28 to (english language and humans and "all adult (19 plus years)") (1189)

Each part was specifically translated for searching the other databases (EMBASE, Web of Science, and Cochrane databases)

Supplementary Material 4. Quality assessment of included risk scores using the PROBAST tool

Author, year of publication	*Risk tool*	*Type or model evaluation*	*Participant selection*	*Predictors*	*Outcome*	*Sample size and participant flow*	*Analysis*	*Overall risk of bias*	*Overall applicability (concern)*	*Usability of model*
Geubbels, 2006	THA specific risk model for SSI	Development	?	+	+	-	-	-	?	No
Mu, 2011	HPRO	Development	?	-	?	-	-	-	?	No
Mu, 2011	KPRO	Development	?	-	?	-	-	-	?	No
Paxton, 2011	NS	Development	?	-	-	-	-	-	-	No
Berbari, 2012	Baseline Mayo PJI risk score	Development	?	-	-	?	-	-	-	No
Berbari, 2012	1-month-postsurgery Mayo PJI risk score	Development	?	-	-	?	-	-	-	No
Bozic, 2013	NS	Development	-	-	-	?	-	-	-	No
Lewallen, 2014	HPRO	External validation	?	-	-	?	?	-	?	Yes
Lewallen, 2014	KPRO	External validation	?	-	-	?	?	-	?	Yes
Inacio, 2015	RxRisk-V for THA	Development	?	-	-	-	-	-	-	No
Inacio, 2015	Elixhauser for THA	Development	?	-	-	-	-	-	-	No
Inacio, 2015	Charlson for THA	Development	?	-	-	-	-	-	-	No
Inacio, 2015	RxRisk-V for TKA	Development	?	-	-	-	-	-	-	No
Inacio, 2015	Elixhauser for TKA	Development	?	-	-	-	-	-	-	No
Inacio, 2015	Charlson for TKA	Development	?	-	-	-	-	-	-	No
Maradit Kremers, 2015	Claims-based risk model for THA	Development	?	-	-	-	-	-	-	No
Maradit Kremers, 2015	Claims-based risk model for TKA	Development	?	-	-	-	-	-	-	No
Tikhilov, 2016	NS	Development	?	-	-	-	-	-	-	No

+	Low risk of bias
?	Unclear risk of bias
-	High risk of bias

HPRO, National Healthcare Safety Network surgical site infections risk model for hip arthroplasty; KPRO, National Healthcare Safety Network surgical site infections risk model for knee arthroplasty; NS, not stated; PJI, periprosthetic joint infection; SSI, surgical site infection; THA, total hip arthroplasty; TKA, total knee arthroplasty